Determination of Dimeric Disulfide Linkage in a Recombinant Human Prolactin Antagonist

A Senior Honors Thesis

Presented in Partial Fulfillment of the Requirements for Graduation with Distinction in Biochemistry in the Undergraduate Colleges of The Ohio State University

by

Mark Troyer

The Ohio State University

June 2005

Project Advisor: Professor Charles L. Brooks, Department of Biochemistry

Abstract

ΔhPRL is a potent human prolactin (hPRL) antagonist, engineered by the deletion of 12 key residues (residues 41 through 52) from wild-type hPRL. The resulting monomeric protein displays potent antagonist activity with little remaining agonist activity, however an undesirable result of this deletion is difficulty in folding, as evidenced by significant formation of dimeric species. The dimeric species is biologically inert. ΔhPRL is derived from wild-type hPRL a 199-residue single peptide protein containing three native disulfide bonds (C4--C11, C58--C174, C191--C199). Using mass spectrometric and additional biochemical techniques, we intend to determine the nature of the linkage found in dimeric ΔhPRL and identify key residues participating in this linkage.

Recombinant ΔhPRL, methionyl protein expressed using BL-21 E. coli, was folded, and purified from cell lysates by ion-exchange chromatography. Monomeric and dimeric species were separated by size-exclusion chromatography. Both species were characterized by absorbance spectroscopy, luminescence spectroscopy, and intact accurate mass determination by electro-spray Q-TOF. Each sample was analyzed by SDS-PAGE under reducing (2-BME) and non-reducing conditions. Dimeric ΔhPRL was denatured in urea, alkylated, and digested with trypsin. Tryptic peptides were analyzed by MALDI-TOF and electrospray LC/MS/MS to generate peptide maps and sequence information.

Accurate mass determination of intact ΔhPRL revealed that the molecular weight of dimeric species is twice the molecular weight of monomeric species, indicating that dimer is made of two constituent monomers with no sequence modification. Comparison

of reducing and non-reducing SDS-PAGE analyses shows that dimeric species present after folding are resolved into purely monomeric constituents when treated with 2-BME, implying that dimeric linkage is due to inter-molecular disulfide bond formation. Further, spectroscopic variation between monomeric and dimeric species indicates that folding is perturbed in the dimeric species, increasing the prospect of opportunistic inter-molecular disulfide formation between cysteine residues. Peptide maps with approximately 70% coverage and sequence data resulting from tryptic digestion of alkylated dimeric ΔhPRL display cysteine-containing peptides in a mixture of disulfide-linked and free (alkylated) states, suggesting random intra-molecular and inter-molecular pairing between available cysteines. Observation of the disulfide-linked peptide 11-16 paired to 11-16 implies C11--C11 disulfide formation. This disulfide can only exist between molecules, suggesting that dimeric linkage is the product of intermolecular disulfide bonding between C11 residues.

Introduction

Human prolactin is a peptide hormone secreted by the anterior pituitary. Prolactin is historically associated with lactation since it acts as a growth factor for mammary epithelium, however it has also been implicated in mammary carcinogenesis (for review, [1]). As a possible target for treatment of breast cancer, the creation of potent prolactin antagonists has been the focus of much research.

Human prolactin is part of a super-family of structurally homologous hormones including human growth hormone and human placental lactogen. These molecules consist of four-helix bundles in an up-up-down-down order such that extensive unstructured regions join the c-terminal ends of helices 1 and 3 to the n-terminal ends of helices 2 and 4, respectively. Native wild-type human prolactin is a 199 amino acid peptide, joined by 3 intramolecular disulfide bonds (Fig 1). One disulfide exists on a small n-terminal loop, between Cys4 and Cys11. A similar disulfide is formed on a small c-terminal loop, between Cys191 and Cys199. A third disulfide is formed in a more central and structurally important region of the molecule, between Cys58 and Cys174.

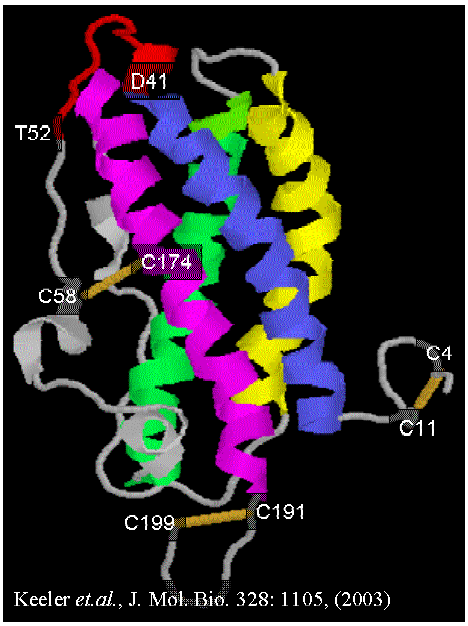


Figure 1: Solution structure of human prolactin; disulfide bonds are shown in gold between C4 and C11, C58 and C174, and C191 and C199.

The disulfide-linked loops on either end of the molecule have little structural influence, but the n-terminal loop in particular has shown a mixed influence on the activity of the molecule [2]. The disulfide formed between Cys58 and Cys174 is commonly referred to as the central disulfide, and

its formation is critical to successful folding and activity of the molecule [3].

The mechanism by which prolactin acts to stimulate growth is key to understanding many approaches of current prolactin antagonists. Human prolactin acts at the cell surface of its target tissue by receptor dimerization (see Fig 2): prolactin is first bound at site 1 by a prolactin receptor; this binding induces a conformation change that opens site 2 so that a second receptor binds prolactin at site 2. The formation of this heterotrimeric complex (prolactin + 2 receptors) is essential to growth stimulation by prolactin. Further, it has been shown that site 2 binding is functionally coupled to site 1 binding, meaning that site 2 binding depends on receptor occupation of site 1 [4].

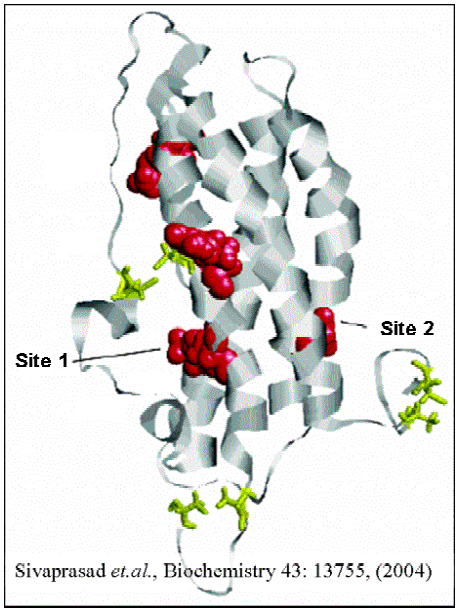


Figure 2: Solution structure of human prolactin; putative binding sites are indicated as site 1 and site 2; disulfide bonds are shown in yellow

Delta prolactin is a potent prolactin antagonist developed by the Brooks lab, engineered by the deletion of 12 key residues from wild-type human prolactin [5]. As shown in Figure 1, residues Asp41 through Thr52 (in red) were removed by site-directed mutagenesis from the region c-terminal to helix 1. This deletion results in a 14,000-fold reduction of activity compared to wild-type prolactin. It is likely that the large loss of activity is due to the functional de-coupling of site 1 and site 2 binding by removal of the coupling residues 41-52.

Although delta prolactin displays promising prolactin antagonism, it also displays marked difficulty in folding, as evidenced by the observation of dimeric and aggregate

forms during isolation. It is therefore worthwhile to determine key residues involved in dimeric delta prolactin linkage, and the nature of their linkage.

Methods

Expression

Expression was carried out using previously described protocols [6,7]. The negative strand of a pT7-7f(-) phagemid – a modification [7] of the pT7-7 vector provided by S. Tabor, Harvard Medical School, Boston, MA – was used for cloning, production of ssDNA, and protein expression in *Escherichia coli* strains DH5α, RZ1032, and BL21(DE3), respectively. Mutagenesis was performed by the Kunkel method [8] from wild-type phagemids containing an n-terminal Met codon. The double-stranded DNA plasmid containing the sequence for delta prolactin was transformed into BL-21(DE3) *E. coli*, which were then plated onto LB-agar plates containing ampicillin. Colonies were selected and grown to in 1 L LB containing ampicillin at 37 ˚C. Upon reaching an approximate OD_{600} of 0.3, expression was induced by addition of IPTG, and cultures were allowed to grow an additional four hours.

Isolation

Following expression, protein was extracted into 100 ml of 100 mM Tris buffer (pH 11.5) containing 4.5 M urea, and allowed to air oxidize for 2 days at 4˚C. The oxidized protein was then dialyzed against 20 mM Tris buffer (4 L, pH 7.5) for six exchanges. During this dialysis, we predict that the protein folds to its native form by first gaining helical structure with the rapid loss of urea, then forming increasingly stable disulfides as the pH drops more gradually toward neutrality. After dialysis, the protein was purified on a DEAE-Sepharose anion-exchange resin in 20 mM Tris buffer (pH 7.5)

using a 0 M to 0.5 M NaCl gradient. Chromatography was monitored at 250 nm and 280 nm, and fractions were collected based on these absorbance measurements. The protein was dialyzed against 10 mM ammonium bicarbonate following chromatography, dried and stored at -30 ˚C.

Size-Exclusion Chromatography

Size exclusion chromatography was used to resolve monomeric and dimeric species. The separation was performed on a Superdex 75 resin in 10 mM ammonium bicarbonate (5 cm diameter, 50 cm length column), and was monitored at 250 nm, 280 nm, and 340 nm.

Absorbance and Luminescence Spectroscopy

All spectroscopic analysis was performed in 10 mM Tris buffer (pH 8.2) containing 150 mM NaCl. For absorbance spectroscopy, 20 μM protein was analyzed using a 240 nm to 320 nm scan on a Perkin Elmer Lambda 45 UV/Vis Spectrometer. For luminescence spectroscopy, the emission spectra of a 1 μM protein sample was analyzed in a Perkin Elmer LS55 Luminescence Spectrometer from 300 nm to 400 nm upon excitation at 285 nm.

SDS-PAGE Analysis

Sodium dodecyl sulfate polyacrylamide gel electrophoresis was performed in a Hoeffer SE600 large format apparatus using a 15% polyacrylamide gel. Samples were boiled in 3% SDS and 5% glycerol under reducing (2% 2-mercaptoethanol) or nonreducing (no 2-mercaptoethanool) conditions.

Intact Mass Analysis

Protein samples were first purified by rp-HPLC on a peptide trap (Microchem BioResources) in 0.1% trifluoroacetic acid using stepwise additions of 0%, 50% and 90% acetonitrile. Fractions containing 50% and 90% acetonitrile were pooled and analyzed at the Ohio State University Campus Chemical Instrument Center Mass Spectrometry and Proteomics Facility using a Micromass Q-TOF II mass spectrometer. Injected sample was resolved by capillary rp-HPLC, then nanospray-ionized for quadrupole time-of-flight analysis. Sample mass was determined to ±2 Da precision.

Digestion and Mass Analysis

Identification of specific residues participating in dimeric linkage was performed using previously determined experimental schemes [8, 9] involving enzymatic digestion followed by MS and tandem MS analysis. Proteins to be digested were solubilized in 10 mM ammonium bicarbonate, then 20 mM iodoacetamide and 2 M urea were added (end concentrations) and the sample was incubated at room temp in darkness for at least 15 minutes to alkylate free Cys residues and denature the protein. Digestion was performed using a 1:25 enzyme to substrate ratio at 37°C for 4 h or 20 h.

The masses of the resulting tryptic peptides were analyzed using matrix assisted laser desorption mass spectrometry and capillary rp-HPLC fed nanospray quadrupole time of flight tandem mass spectrometry on a Bruker Reflex III MALDI-TOF and a Micromass Q-TOF II, respectively. MALDI generated peak lists and MS/MS generated fragment ion data were analyzed using in house software (OSU CCIC MS and Proteomics Facility) or online software such as MASCOT [10] or Protein Prospector [11] to deduce sequence information.

Results and Discussion

Protein Folding Studies

Absorbance and luminescence spectra for monomeric wild-type prolactin, monomeric delta prolactin, and dimeric delta prolactin are compared in Figs 3 and 4. Each series is the average of three or more trials, and is normalized at 280nm and 340nm for absorbance and luminescence, respectively. Both spectra are good qualitative measures of the condition of a protein's aromatic residues (Trp, Tyr, Phe) and in some cases a protein's disulfide bond condition; both measures are useful folding indicators.

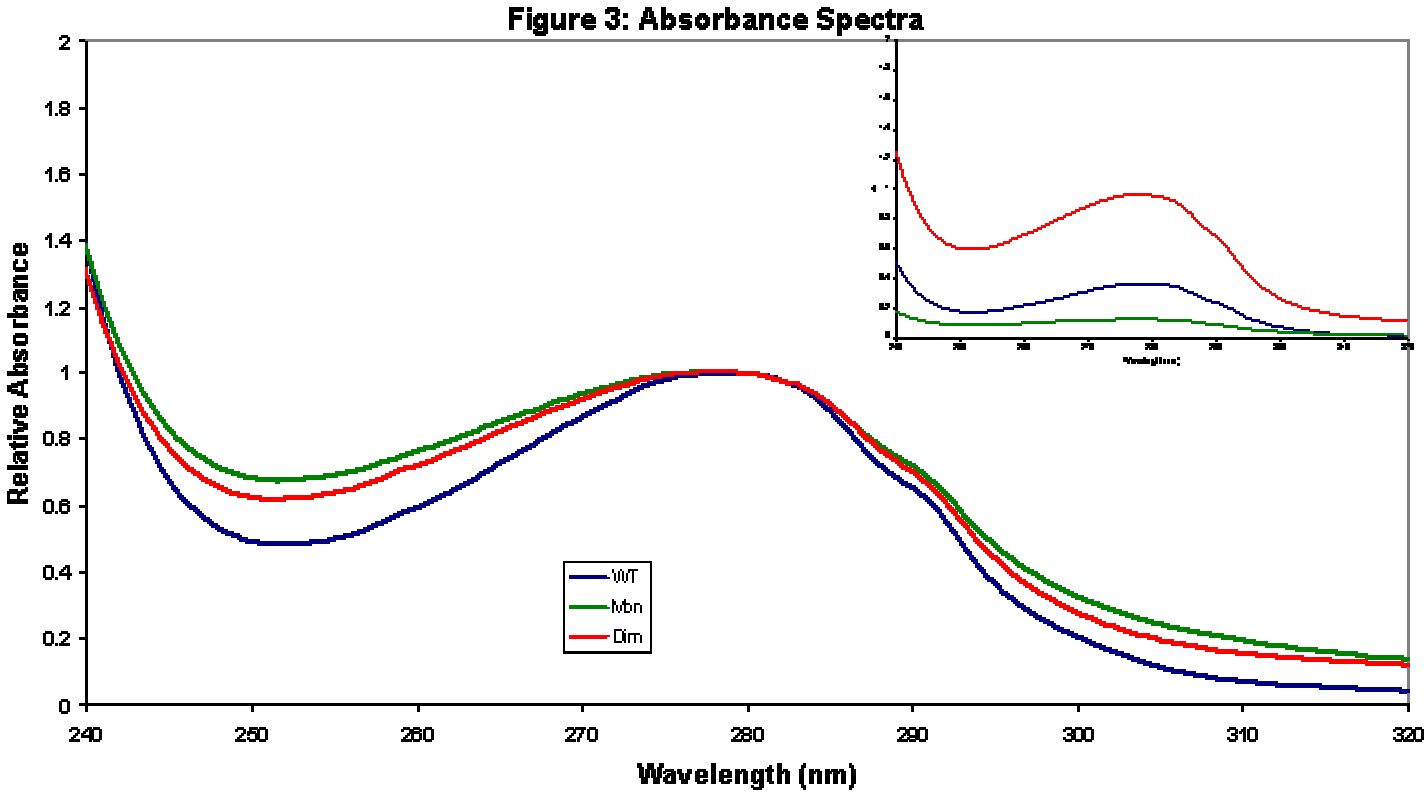


Figure 3: Absorbance spectra for monomeric wild-type (WT, blue), monomeric delta (Mon, green), and dimeric delta (Dim, red) prolactin; spectra are normalized at 280nm, with raw spectra displayed in the inset.

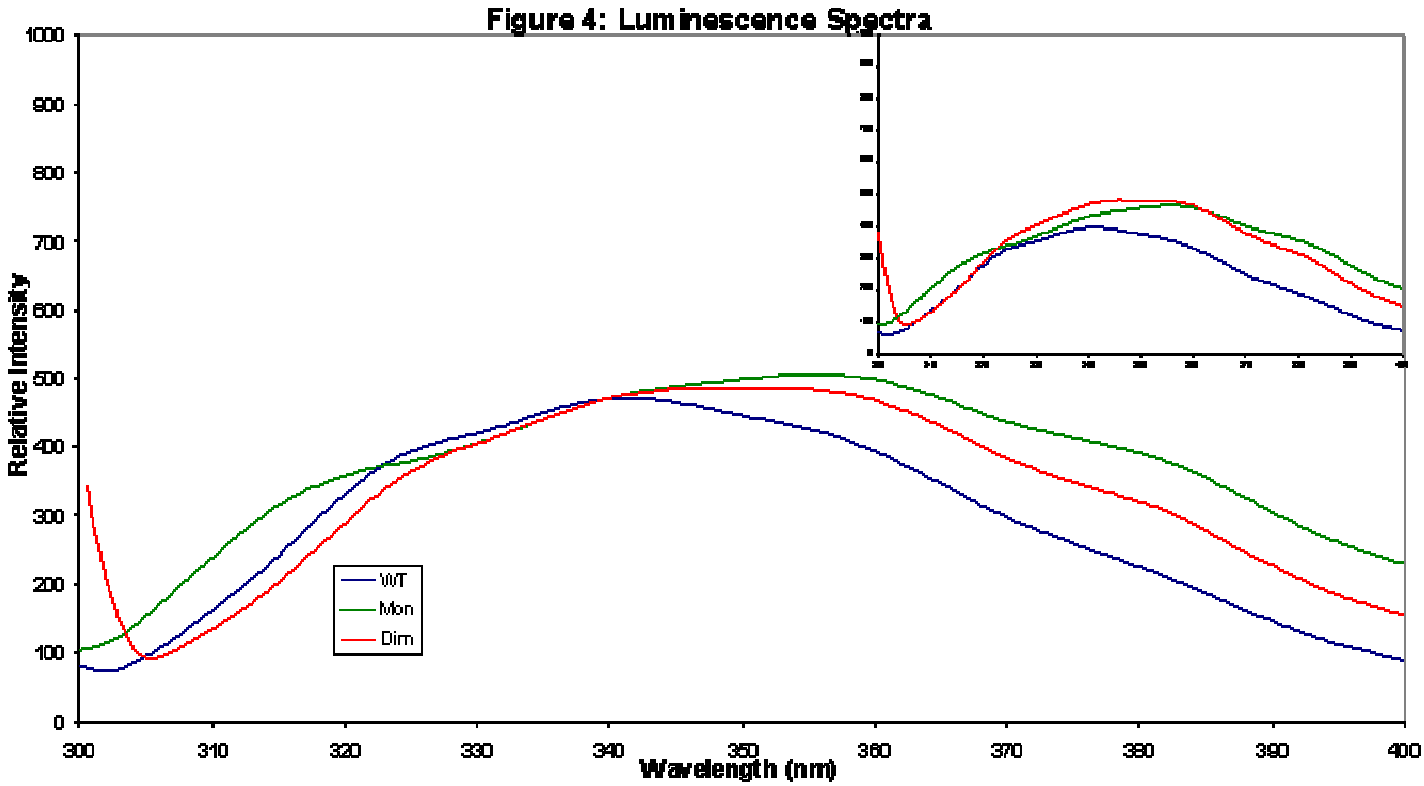


Figure 4: Luminescence spectra for monomeric wild-type (WT, blue), monomeric delta (Mon, green), and dimeric delta (Dim, red) prolactin; spectra are normalized at 340nm, with raw spectra displayed in the inset.

By analyzing a prolactin mutant's 280nm / 250nm absorbance ratio, it is believed that one can deduce how well-folded the molecule is. Interestingly, dimeric delta prolactin displays a ratio that is intermediate between monomeric delta prolactin and monomeric wild-type prolactin. Dimeric delta prolactin was expected to yield 280nm / 250nm ratio lower than monomeric delta prolactin; however this observation may actually indicate that the stress on the central disulfide often present in monomeric delta prolactin is no longer present, perhaps due to alternate disulfide formation.

Comparison of luminescence spectra between the aforementioned molecules was equally puzzling. Analysis of luminescence spectra is particularly useful for indicating the Trp residue environment: more polar environments stabilize excited Trp residues, thus a red shift in luminescence spectra is often indicative of Trp residues exposed to an

aqueous environment due to perturbed folding. Dimeric delta prolactin displays a distinct red shift when compared to wild-type prolactin, indicating relatively perturbed folding. Monomeric delta prolactin displays more red shift than dimeric prolactin, in addition to a unique increase in luminescence at 315nm, an area more indicative of Tyr condition. While Tyr residues are often quenched by nearby Trp residues, this increase may be due to fluorescence of a Tyr no longer nearby Trp residues.

Both dimeric and monomeric delta prolactin display absorbance and luminescence characteristics consistent with stressed structure. One can surmise that the 12 residue deletion responsible for delta prolactin's potent antagonism also puts intense stress on the molecule, and decreases the likelihood of properly folded molecules with properly paired disulfides.

The Nature of Dimeric Delta Prolactin Linkage

Comparison of reducing and nonreducing SDS-PAGE is a robust indicator of protein purity and disulfide structure. For the purposes of this study, such a comparison is an excellent indicator of dimeric linkage type (Fig 5). 2-Mercaptoethanol is used to reduce proteins, essentially competing for constituent Cys residues of a disulfide bond. This results in two Cys residues each bound to 2-mercapto-ethanol rather than to each other.

Three samples of dimeric delta prolactin were analyzed by nonreducing and reducing SDS-PAGE. While samples in nonreducing conditions appear predominantly dimeric (~90%), the same samples appear predominantly monomeric (~90%) in reducing conditions. This is a very strong indicator that the linkage found in dimeric delta prolactin

is an intermolecular disulfide bond, but it does not indicate between which residues this bond is formed.

Intact mass analysis was performed on a mixture of purified monomeric and dimeric delta prolactin (Fig 6), using mass spectrometry. This analysis shows with great precision (± 2 Da < 0.005% error) the masses of monomeric and dimeric delta prolactin, indicating that dimeric delta prolactin is composed solely of unmodified monomeric constituents. This data further implicates an intermolecular disulfide dimeric linkage.

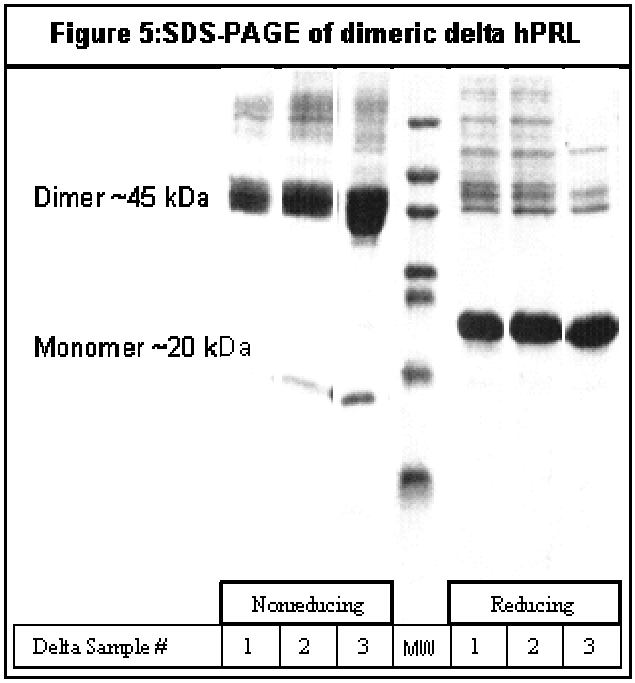
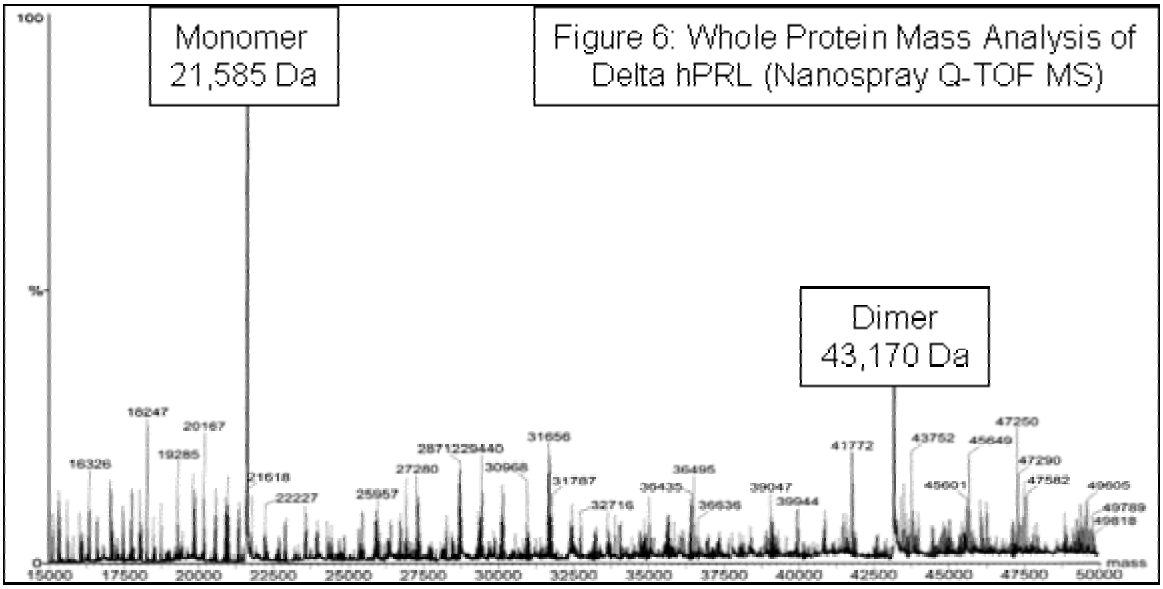


Figure 5: SDS-PAGE of dimeric delta prolactin samples under reducing and nonreducing conditions. Samples are indicated at the bottom of their corresponding lanes, with their conditions indicated above the sample number.

Figure 6: Mass analysis of whole protein mixture of purified monomeric and dimeric delta prolactin; theoretical mass, calculated using sequence data, is 21,583 Da per monomer; nanospray Q-TOF MS is precise to ± 2 Da.



Key Residues Participating Dimeric Delta Prolactin Linkage

An experimental method was prepared using previous protocols in order to identify key residues participating in the intermolecular disulfide linkage found in dimeric delta prolactin. Digestion with the endopeptidase trypsin cleaves proteins at specific residues (Lys and Arg) resulting in a mixture of peptides of predicted sequence and mass. Tryptic digestion has particular utility in disulfide analysis. Tryptic digestion using nonreducing conditions can isolate disulfide-linked peptides – two tryptic peptides bound together by a disulfide bond. These are useful, because they allow us to identify which Cys residues are bound to each other by disulfide bonds.

MALDI-TOF mass spectrometry yields the same type of mass analysis as nanospray Q-TOF, and differs only in the ionization method. MALDI is used is this case, because it is less sensitive to the presence of undesired ions (ie salts, buffers). Although observation of a peak of matching mass to a certain theoretical digest product is a strong indication of that product's presence. Tandem mass spectrometry can be used to yield much more convincing results.

Tandem mass spectrometry is performed on the Micromass Q-TOF II, but is more complex than the previously mentioned intact mass analysis. During tandem mass spectrometry, ionized sample molecules are first separated by mass (as before), then introduced into a collision cell where the stream of sample molecules is met with a perpendicular stream of inert argon. During this collision with argon, peptide samples fragment predictably at peptide bonds between the amino acid residues, from either n- or c- termini toward the center. Since this fragmentation pattern can be predicted, data obtained from tandem mass spectrometry can be compared to theoretical fragmentation

patterns to yield sequence information. By analyzing the results of a tryptic digestion using mass spectrometry and tandem mass spectrometry, one can gain precise sequence details regarding which residues are responsible for the disulfide linkage in dimeric delta prolactin.

Comparative analysis of monomeric and dimeric delta prolactin using the above experimental scheme yielded >80% sequence coverage for both samples. Since Cys containing peptides were the focus of our study, it is more important to determine which peptides were observed and which were not; all Cys containing peptides were observed in either unalkylated, alkylated, or disulfide linked form except peptide #1 (residues 1-10) containing Cys4. Careful analysis of tandem mass spectrometry fragment ion patterns in dimeric delta prolactin yielded a pattern with the following sequence: Arg-Leu-Thr-Val-Gln-Cys-Cys-Gln (spectrum shown in Fig 7). This sequence is part of the theoretical

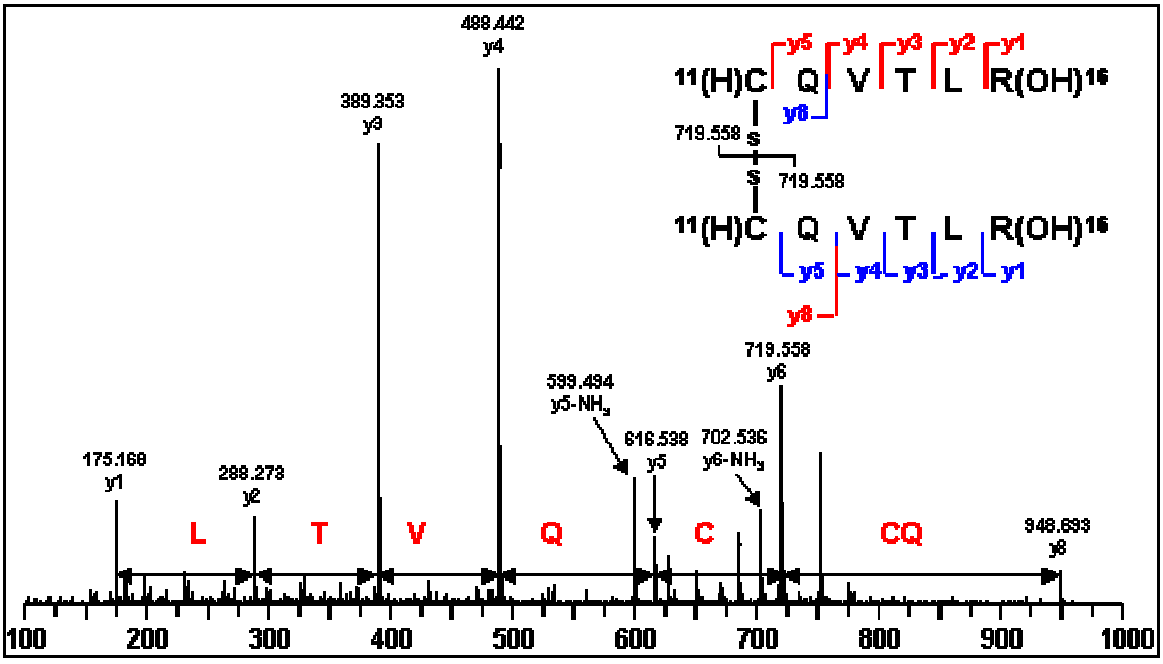


Figure 7: Tandem Mass Spectrum of Cys11--Cys11 linked peptide found only in dimeric delta prolactin; Mass+H of linked peptide=1435.77 Da; Mass+H of precursor ion = 479.27*3 = 1435.81 Da.

tryptic peptide that results from Cys11--Cys11 linkage. The precursor ion had a mass/charge ratio of 479.27, which is consistent with the triply charged ion of the Cys11--Cys11 theoretical peptide (mass 1435.8 Da / charge 3+). Neither this sequence nor its precursor ion mass were observed in monomeric delta prolactin.

From this sequence and precursor ion data, it is reasonable to assume that a disulfide bond exists between two Cys11 residues on different molecules, as one molecule of delta prolactin can only contain one Cys11. Therefore the linkage found in dimeric delta prolactin is likely an intermolecular disulfide bond between Cys11 residues.

Final Conclusions and Future Research

This insight into the nature of and key residues involved in dimeric linkage in delta prolactin allows us to now target intermolecular Cys11--Cys11 disulfide formation, and design new prolactin mutants that have improved folding characteristics. One possibility for new design would be truncation of the n-terminus, excluding the first 11 residues, and removing Cys11 from the molecule entirely [2]. Such a molecule may be an ideal therapeutic candidate, exhibiting both potent prolactin antagonism and robust folding characteristics.

Acknowledgements

I am grateful to Laura Depalatis and Jeffery Voorhees of the Brooks Lab for their support and guidance, to Rhonda Pitsch and Ben Jones of CCIC MSP for their assistance, and to Liwen Zhang and Kari Green-Church of CCIC MSP for their collaboration. Special thanks are due to Dr. Fred Hitzhusen and Dr. Richard Swenson for their participation in my thesis committee, and to Dr. Charles Brooks for his motivation, instruction, and supervision.

References

1. Clevenger CV, Furth PA, Hankinson SE, Schuler LA. Endocrine Reviews **24**(1): 1-27, 2003.
2. Bernichten S, Jomain JB, Kelly PA, Goffin V. Molecular and Cellular Endocrinology **208**(1-2): 11-21, 2003.
3. Luck DN, Gout PW, Sutherland ER, Fox K, Huyer M, Smith M**. Protein Engineering 5**(6): 559-567, 1992.
4. Sivaprasad U, Canfield JM, Brooks CL. Biochemistry **43**: 13755-13765, 2004.
5. Peterson FC, Brooks CL. Protein Engineering, Design & Selection **17**(5): 417-424, 2004.
6. Maciejewski PM, Peterson FC, Anderson PJ, Brooks CL. J. Biological Chemistry **270**: 27661-27665, 1995.
7. Peterson FC, Anderson PJ, Berliner LJ, Brooks CL. Protein Expression and Purification **15**: 16-23, 1999.
8. Stone KL, Williams KR. in The Protein Protocols Handbook Walker JM (ed). Humana Press: Totowa, NJ. 415-425, 1996.
9. Yen TY, Joshi RK, Yan H, Seto NOL, Palcic MM, Macher BA**.** J. Mass Spectrometry **35**:990-1002, 2000.
10. Perkins DN, Pappin DJ, Creasy DM, Cottrell JS. Electrophoresis **20**(18): 3551-3567, 1999.
11. Clauser KR, Baker PR, Burlingame AL. Analytical Chemistry **71**(14): 2871-2882, 1999.